NFT Security: Rug Pull Detection

SW2

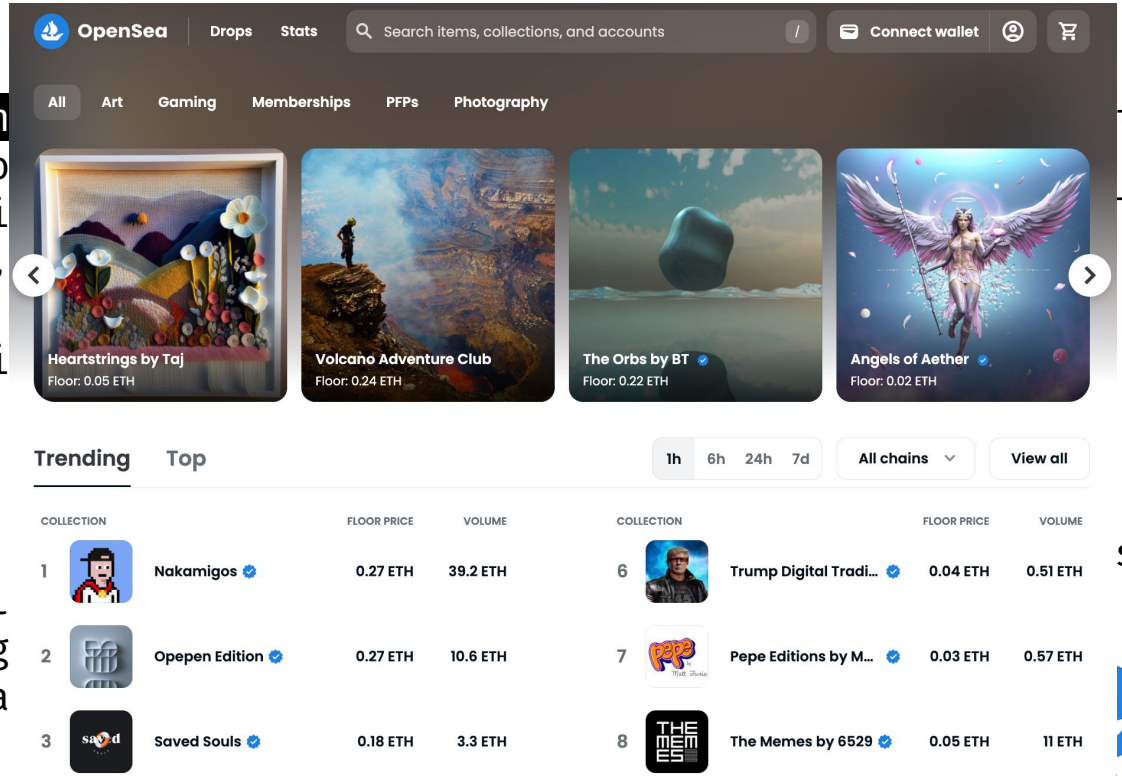
Qi WU: 20677500
Supervised by Prof. Shuai WANG, second reader Prof. Qi Feng CHEN
Attributes to Mr. Zonjie LEE

Table of Contents

01 Introduction

As an important form of cryptocurrency, non-fungible tokens (NFTs) has drawn much public attention. This study focuses on using a supervised model to detect a specific type of NFT fraud: Rug-Pull.

NFT (Non
uses blo
authenti
content,
enabled
populari



OpenSea
and coll
the larg
blockcha

OpenSea

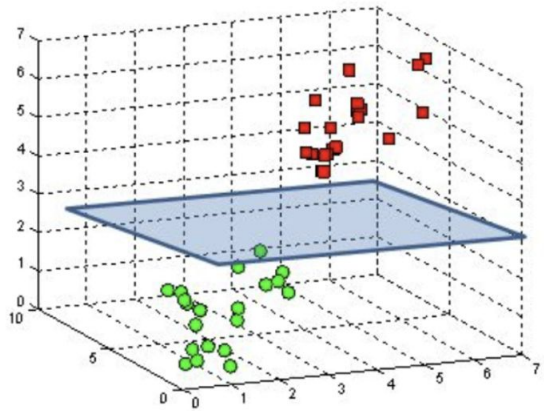
Rug pull is a common type of scam that happens where a project or token creator take investors‘ money and abandon their projects, leaving investors with worthless holdings.

Even OpenSea has experienced a significant number of Rug-Pull scams. In the most famous Rug Pull, *Frosties*, $1.3m US dollars were stolen.

Frosties

SVM (Support Vector Machine) is a machine learning algorithm used for classification and regression analysis. It works by finding the optimal hyperplane that maximizes the margin between different classes of data.

In this project, 2 versions of SVM models have been trained.



Data collection

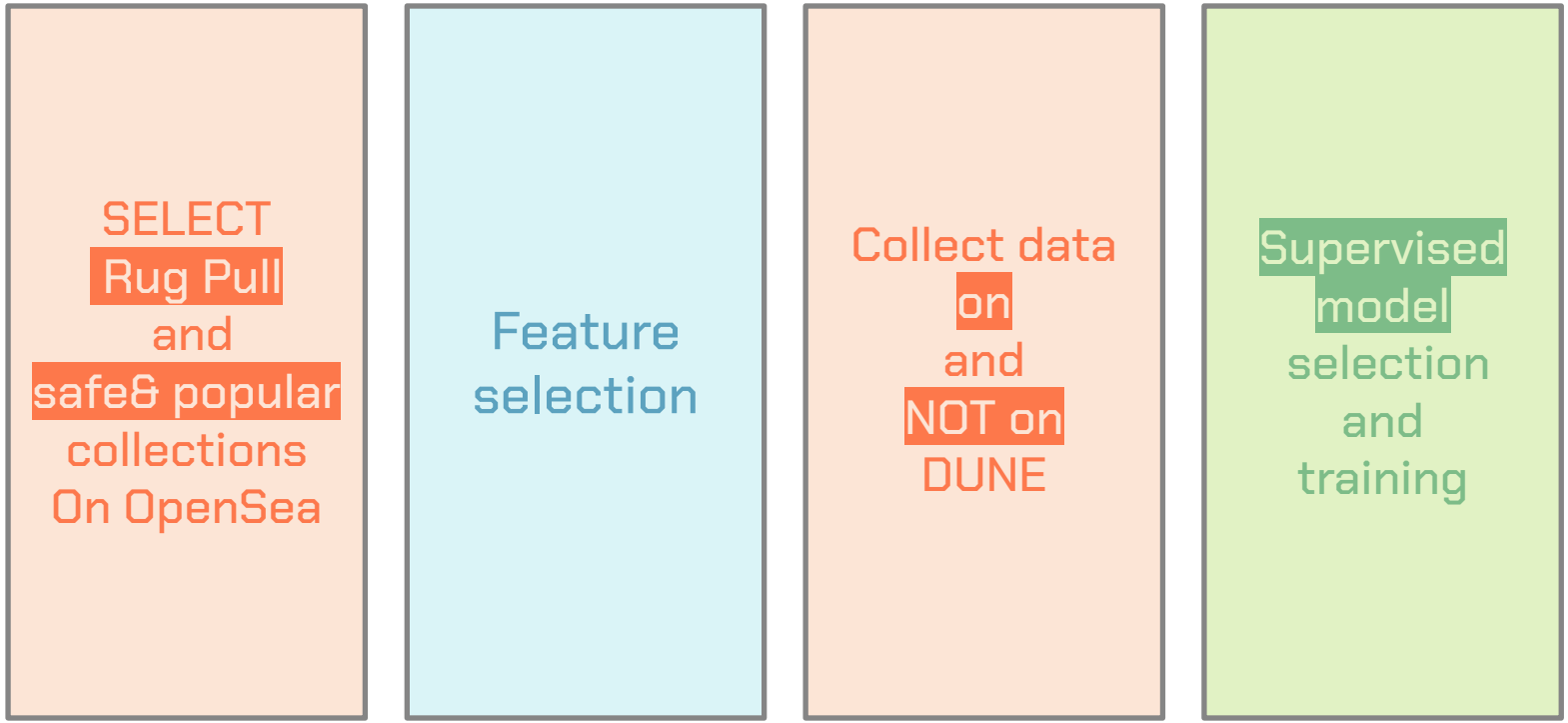
DUNE [SQL]	Select popular OpenSea collections.
Rug Pull Finder	Create a list to collect Rug Pulls.
Twitter	Collection related information (Anonymity, etc.)
OpenSea	Basic information of a collection.

Program Implementation

Colab	Google online coding platform
Python Libraries	-- *matplotlib, numpy, panda, drive, sklearn*
Google Sheet	An online spread sheet tool

How was this project designed and implemented?

02 Workflow


SELECT
Rug Pull
and
safe& popular
collections
On OpenSea
Feature
selection
Collect data
on
and
NOT on
DUNE
Supervised
model
selection
and
training
Data collection I
Feature Engineering
Data collection II
Supervised Learning

SELECT popular collections on OpenSea

To identify the most popular collections on OpenSea, the project uses SQL query language to analyse on-chain data decoded by DUNE.

Popular collections are identified to have high total volumes and potentially more influential.

	COLLECTION	rug pull	owners(#)	items(#)
2	Bored Ape Yacht Club	-1	5,693	9,998
3	Mutant Ape Yacht Club	-1	11,218	19,465
4	CryptoPunks	-1	3,799	9,998
5	Bored Ape Kennel Club	-1	5,438	9,602
6	HV-MTL	-1	10,009	26,398
7	Pudgy Penguins	-1	4,616	8,892
8	CLONE X - X TAKASHI MURA	-1	9,634	19,471
9	Nakamigos	-1	5,674	19,894
10	Opepen Edition	-1	4,944	15,985
11	Otherdeed for Otherside	-1	19,753	61,776
12	Something Official	-1	7,415	19,950
13	The Captainz	-1	4,639	9,999
14	Milady Maker	-1	3,311	9,978
15	NFT Worlds	-1	864	9,999
16	DeGods	-1	836	8,500
17	mfers	-1	5,583	10,019
18	Otherside Koda	-1	3,061	4,996
19	CyberBrokers	-1	3,135	10,000
20	Otherdeed Expanded	-1	11,900	37,568
21	Moonbirds	-1	6,478	10,000
22	MeteoriaNFT	-1	1,925	5,998

```
(SELECT
    ROW_NUMBER() OVER (ORDER BY SUM(price_usd) DESC) AS "Rank",
    CASE
        WHEN art.slug IS NOT NULL THEN CONCAT('<a href="https://dune.com/cryptuschrist/art-blocks-custom-dashboard?Project+ID=',RO
        ELSE COALESCE(collection_opensea_dune,CONCAT('<a href="https://dune.xyz/cryptuschrist/CUSTOM-NFT-Dashboard?NFT%20Contract=
    END AS collection,
    COUNT(*) AS sales,
    SUM(price_usd) AS usd_volume,
    SUM(price) AS total_volume,
    ROUND(MIN(price::numeric),5) AS lowest_sale,
    percentile_cont(.5) within GROUP (ORDER BY price) AS median,
    MAX(price) AS high_sale,
    COALESCE(art_blocks_owners.owners,s.owners) AS "Owners",
    COALESCE(art_blocks_owners.items,s.items) AS "Items",
    fee/10000 AS "CREATER FEE %",
    dune_user_generated.gemlink(a.nft_contract_address) || dune_user_generated.bytea2etherscantoken(a.nft_contract_address) AS lin
FROM dune_user_generated."cryptuschrist_allexchange3" a

LEFT JOIN (SELECT collection,  CONCAT('https://opensea.io/collection/', slug) AS url, slug, token_id AS art_token, nft_contract_ad
LEFT JOIN (SELECT project_id, owners, items FROM dune_user_generated.cryptus_art_blocks_owners_items) art_blocks_owners on art_blo
LEFT JOIN dune_user_generated."cryptuschrist_nft_projects5" on "nft_contract" = a.nft_contract_address
```

DUNE

Save and run

Find Rug Pull collections – *Rug Pull Finder & etc.*

This project builds its database using the list of confirmed rugs reported by Rug Pull Finder and revalidate the Rug on creditable news platforms whenever relevant information is available.

To increase the model's robustness, Rug Pulls selected have deliberately been made to have different properties.

1	COLLECTION	rug pull	owners(#)	items(#)
79	SHIBABEAST	1	1401	2515
80	The-Companion	1	4944	8884
81	Swipa The Fox	1	-1	-1
82	Rowdy Society RSNFT	1	411	770
83	Rich Bulls Club	1	2121	3495
84	the animoon	1	-1	-1

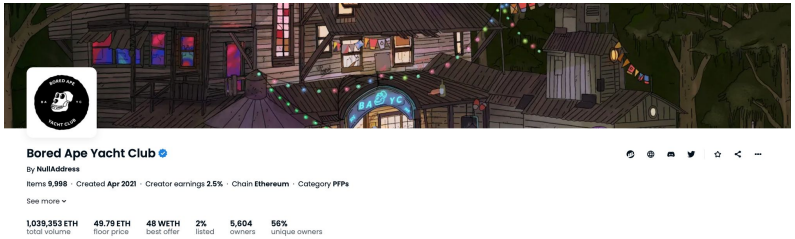
RUG PULL FINDER

Data Collection

- via DUNE and manually

Although much data has been made available via DUNE, a few Rug-Pull-relevant information of an NFT collection are manually collected.

These data is only available through a collection's website and social media accounts, for whom, more comprehensive analysis is required.

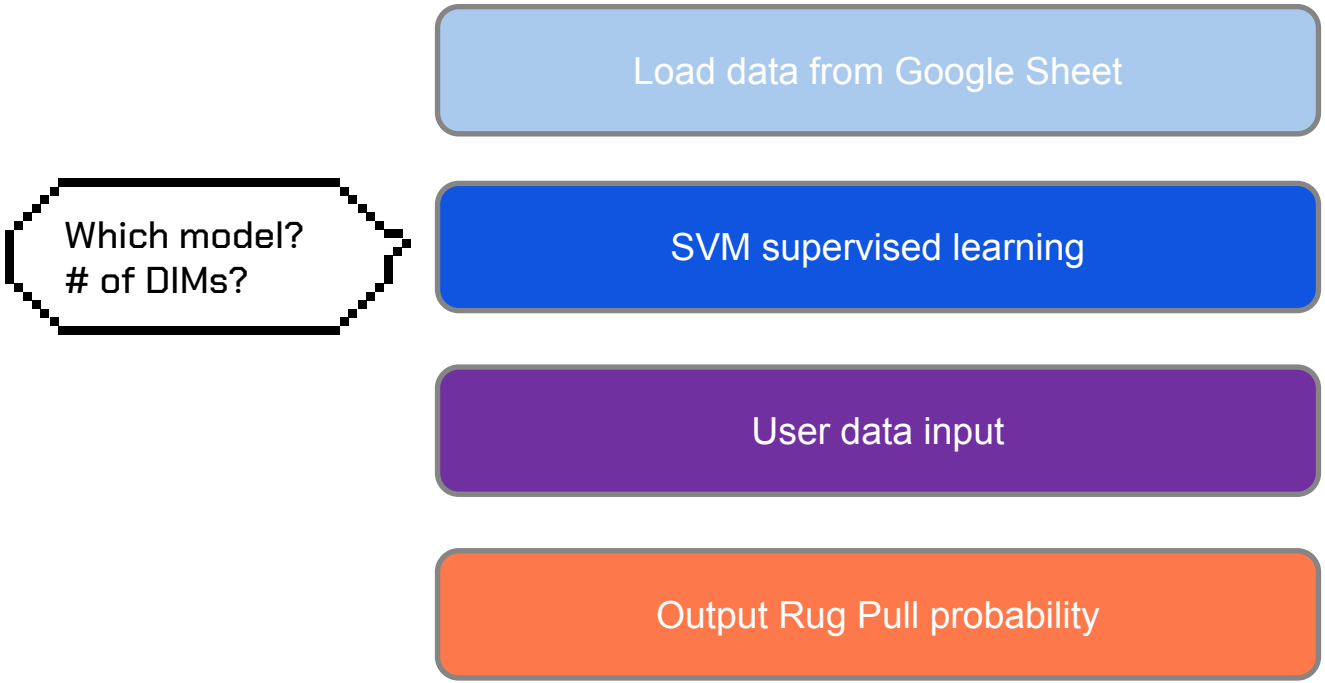


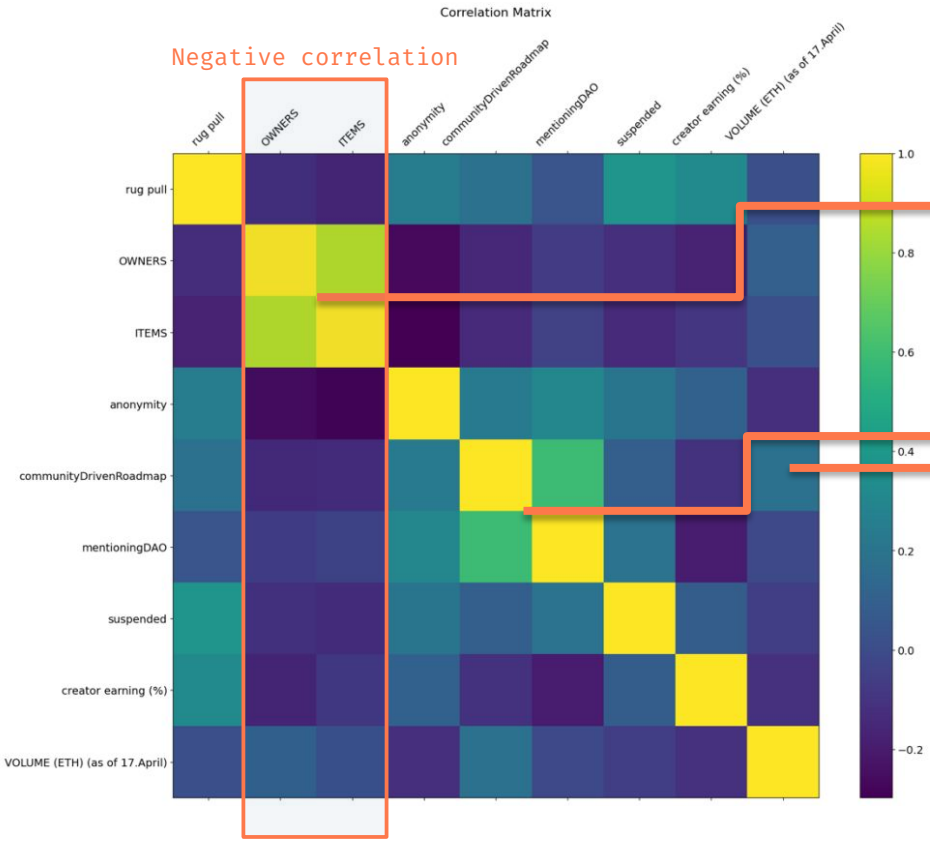
Data conveniently available on DUNE	Creator Earning Total Volume Number of Items Number of Owners
Information manually collected	Anonymity Community Driven Roadmap Mentioning DAO Suspension on OpenSea If_Rug Pull

Group	Item	Description
creator information	anonymity	Whether the NFT creators are hiding their real identities while participating in the market.
community reliance	community Driven Roadmap	Whether the collection is involving its community in the development and direction of the project's roadmap or future plans.
	Mentioning DAO	Whether the NFT collection is acknowledging the use of decentralized autonomous organizations (DAOs) in their governance or decision-making processes.
OpenSea metadata	suspension	Whether the collection is currently removed from a marketplace due to a violation of the platform's rules or terms of service.
	creator earning (%)	The amount of revenue or profit that the creator or creators of a collection earn from the sales of its NFTs on a marketplace.
	Items	the total count of NFTs that have been minted or created and made available for sale within a particular collection on OpenSea
Trading data	owners	the count of unique wallet addresses that hold at least one NFT from a particular collection on a marketplace or blockchain.
	volume (ETH)	Sum of all sales or transactions that have occurred for NFTs within a particular collection on OpenSea. It represents the total value exchanged for the NFTs within that collection.
Rug pull	Rug pull	whether this particular collection or NFT has been confirmed to be a fraudulent project known as Rug Pull

Decentralized Autonomous Organization

The reliability of an NFT project is also relevant to their Social engineering status.


Load data from Google Sheet
Which model?
# of DIMs?
SVM supervised learning
User data input
Output Rug Pull probability



Number of Owner
&
Number of Items

Community Driven Roadmap
&
Mentioning DAO

Total volume
&
Community Driven Roadmap

Creator Earning
Anonymity
Community Driven Roadmap
Mentioning DAO
Suspension on OpenSea
If_Rug Pull

SVM

5 DIM vs 8 DIM

Creator Earning
Anonymity
Community Driven Roadmap
Mentioning DAO
Suspension on OpenSea
If_Rug Pull
Total Volume
Number of Items
Number of Owners

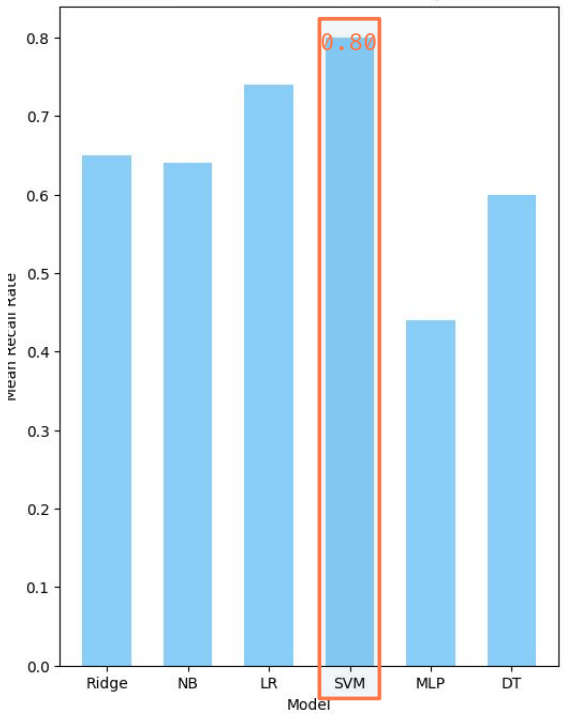
```
[ ] from sklearn.svm import SVC
    svm = SVC(kernel='linear', C=1, random_state=42)

    svm.fit(X_train, y_train.values.ravel())
    svm.predict(X_test), svm.score(X_test, y_test)
```

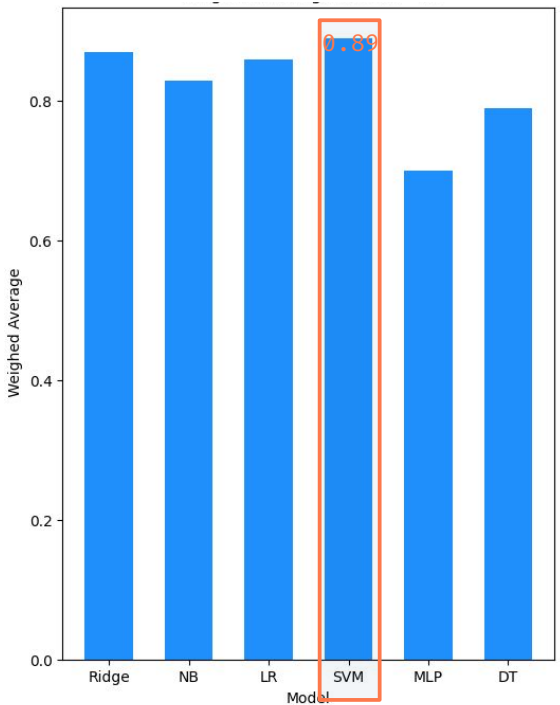
```
(array([-1, -1,  1, -1, -1, -1, -1, -1, -1, -1,  1, -1, -1, -1,  1, -1, -1,
        -1, -1]),
 0.8947368421052632)
```

scikit learn

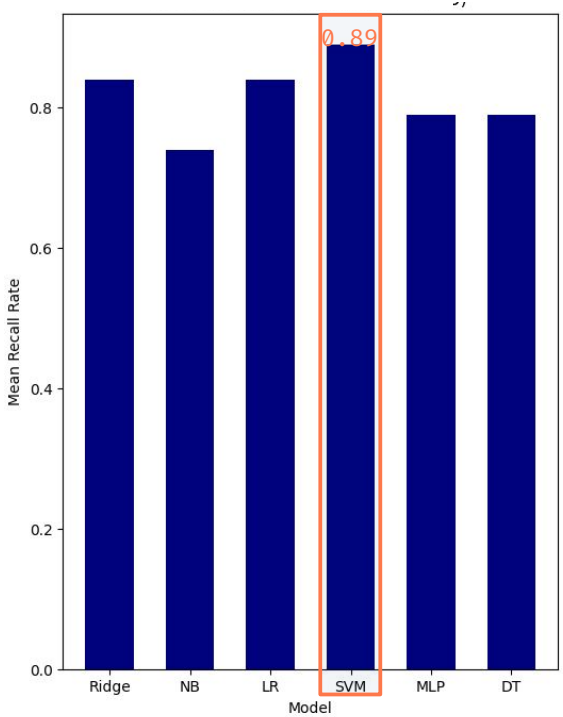
Macro average
F1 score



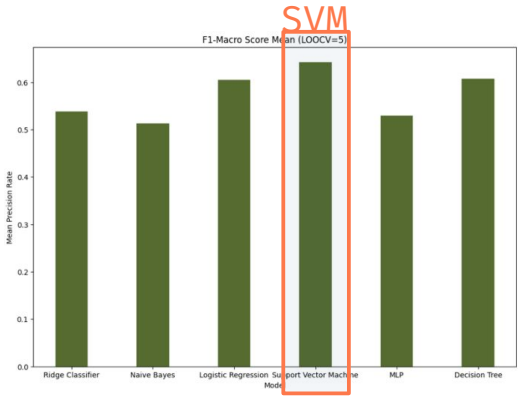
Weighted Average
precision

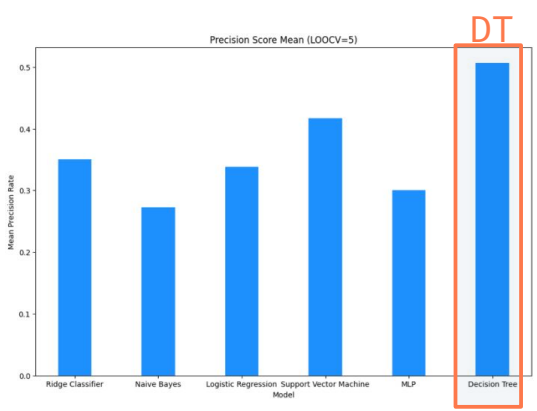


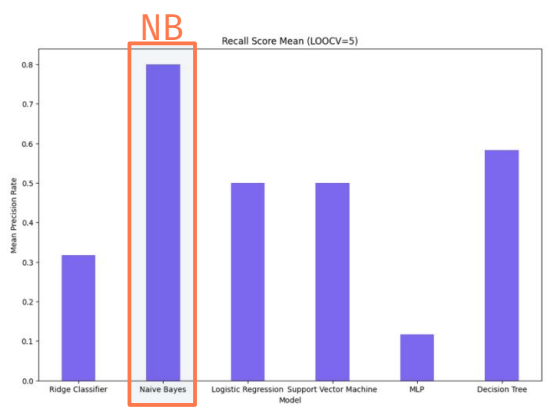
Weighted Average
recall



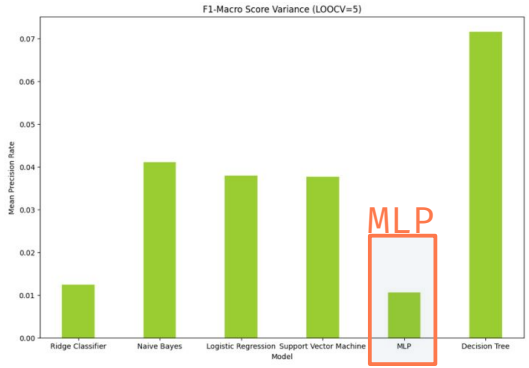
MEAN

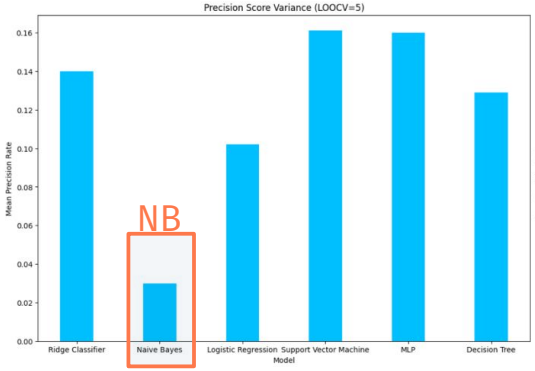

SVM
F1-Macro Score Mean (LOOCV=5)
Mean Precision Rate
Ridge Classifier
Naive Bayes
Logistic Regression
Support Vector Machine
MLP
Decision Tree

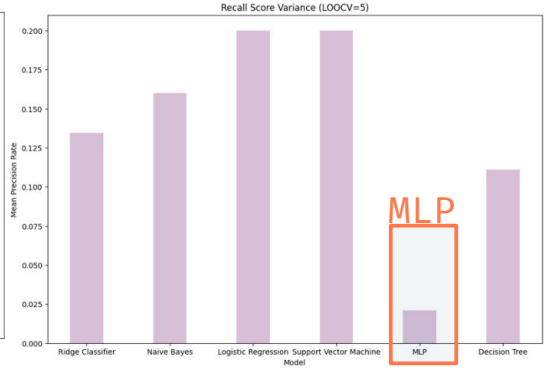

DT
Precision Score Mean (LOOCV=5)
Mean Precision Rate
Ridge Classifier
Naive Bayes
Logistic Regression
Support Vector Machine
MLP
Decision Tree
Model

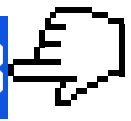

NB
Recall Score Mean (LOOCV=5)
Mean Precision Rate
Ridge Classifier
Naive Bayes
Logistic Regression
Support Vector Machine
MLP
Decision Tree
Model

VAR

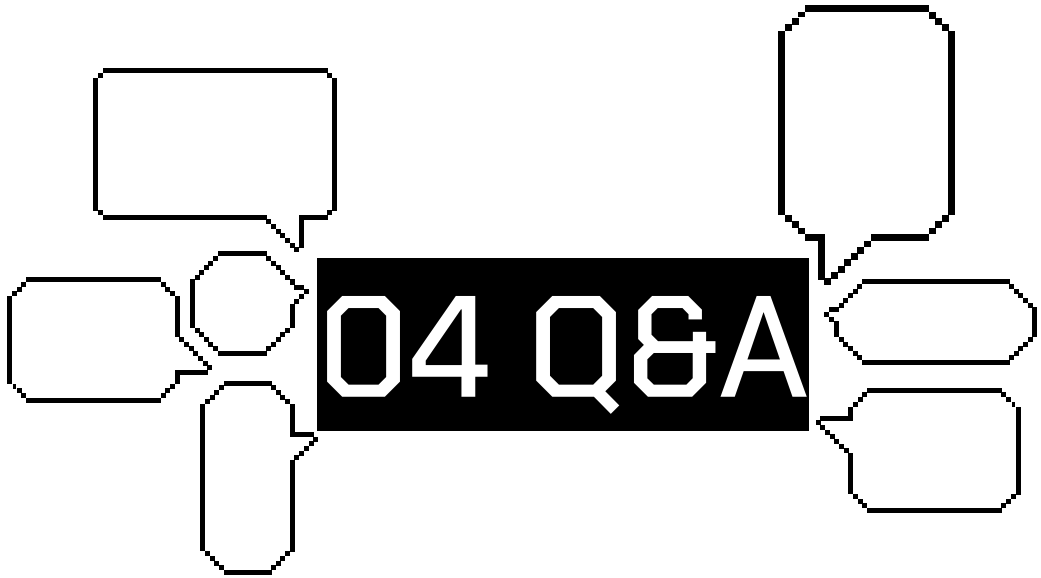

F1-Macro Score Variance (LOOCV=5)
MLP
Mean Precision Rate
Ridge Classifier
Naive Bayes
Logistic Regression
Support Vector Machine
MLP
Decision Tree
Model

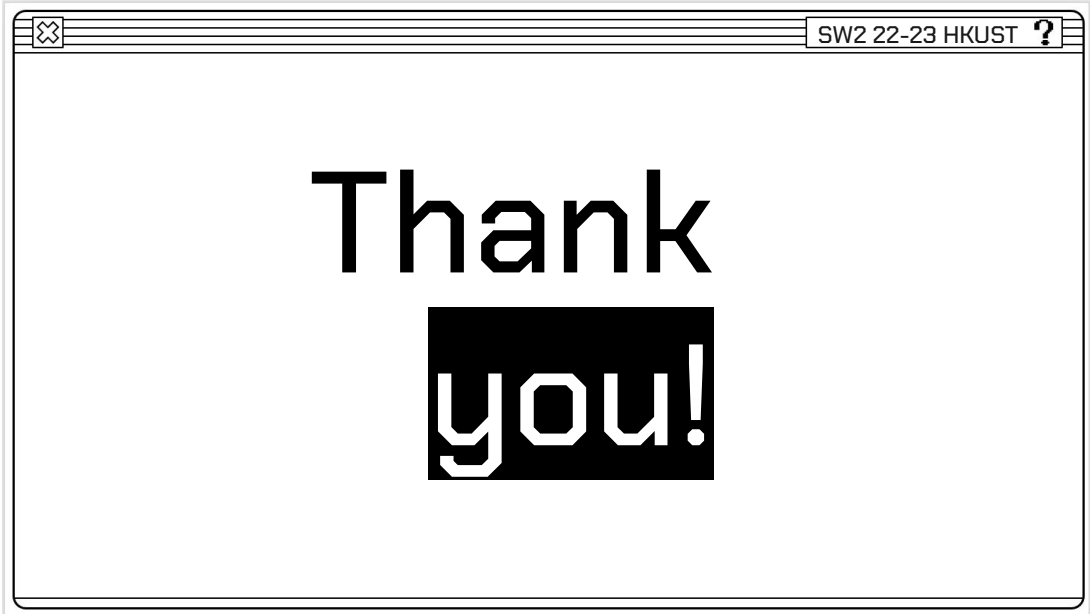

Precision Score Variance (LOOCV=5)
NB
Mean Precision Rate
Ridge Classifier
Naive Bayes
Logistic Regression
Support Vector Machine
MLP
Decision Tree
Model


Recall Score Variance (LOOCV=5)
MLP
Mean Precision Rate
Ridge Classifier
Naive Bayes
Logistic Regression
Support Vector Machine
MLP
Decision Tree
Model



03 Demo




SW2 22-23 HKUST
Thank
you!